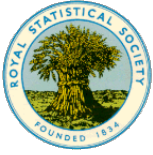


Statistics for Medical Research (Part II)

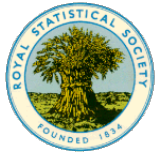
***Normal Distribution
Parametric Tests
Non-Parametric Tests***

Dr. Wong Kai Choi

***Department of Psychiatry
Pamela Youde Nethersole Eastern Hospital
23rd February 2016***



The Normal Distribution

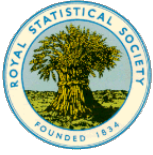


The Normal Distribution

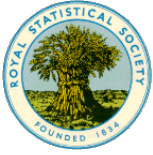
正態分佈

The Gaussian Distribution

高氏分佈

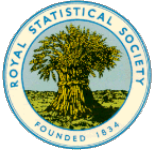


Johann Carl Friedrich Gauss 高斯 (1777 – 1855)



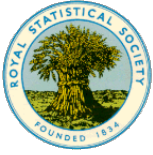
Why we learn Normal distribution?

- It is often appropriate to use the normal distribution as the distribution of a population or random sample.
- The normal distribution is frequently employed in inferential statistics.



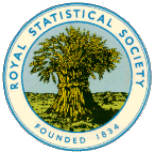
Probability density function of normal distribution

$$f(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

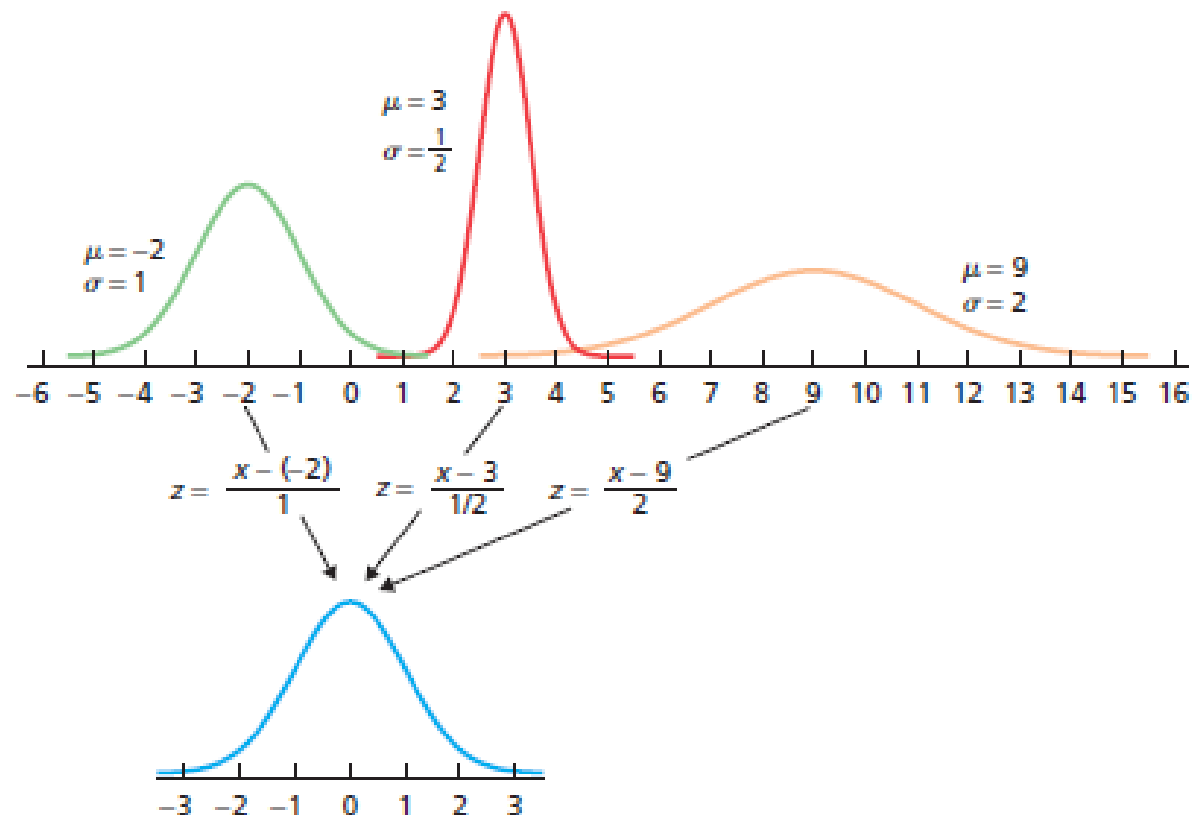


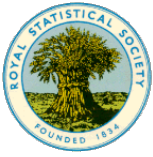
Probability density function of standard normal distribution

$$\phi(x) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$$

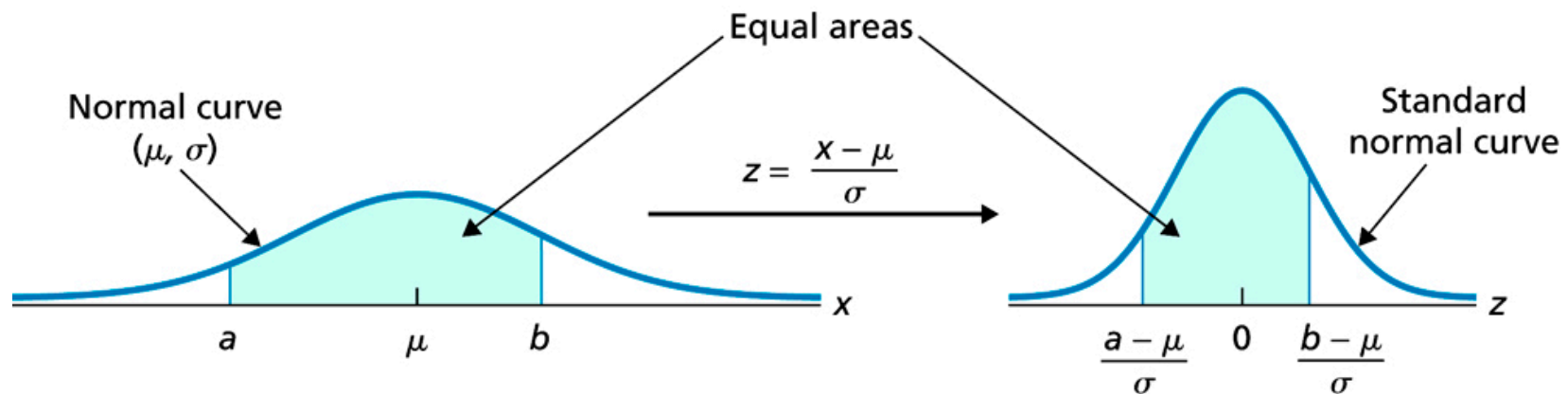


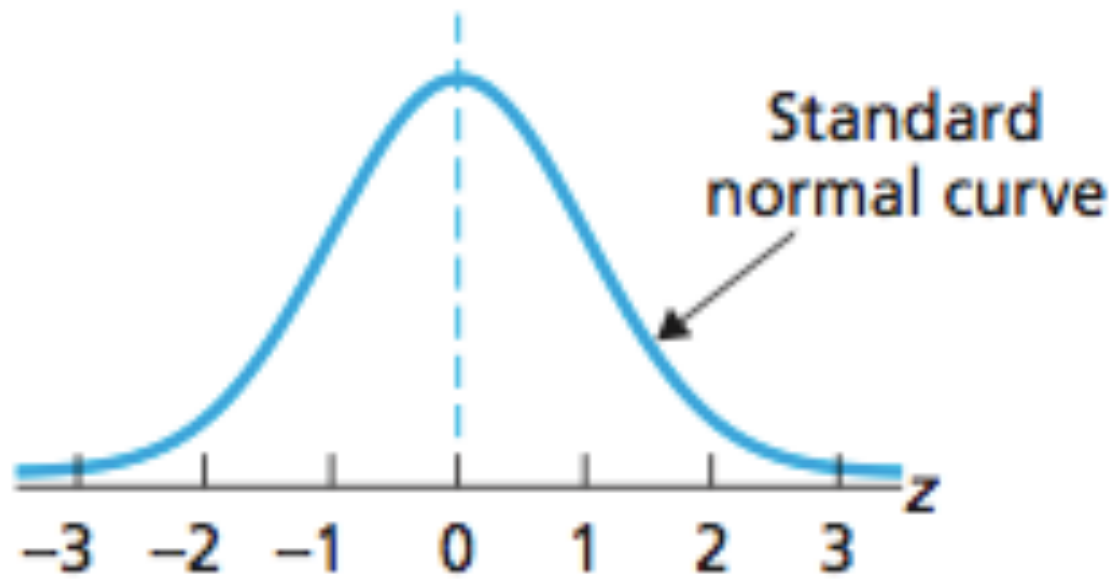
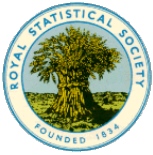
Standardizing normal distributions

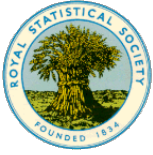




Finding percentages for a normally distributed variable from areas under the standard normal curve







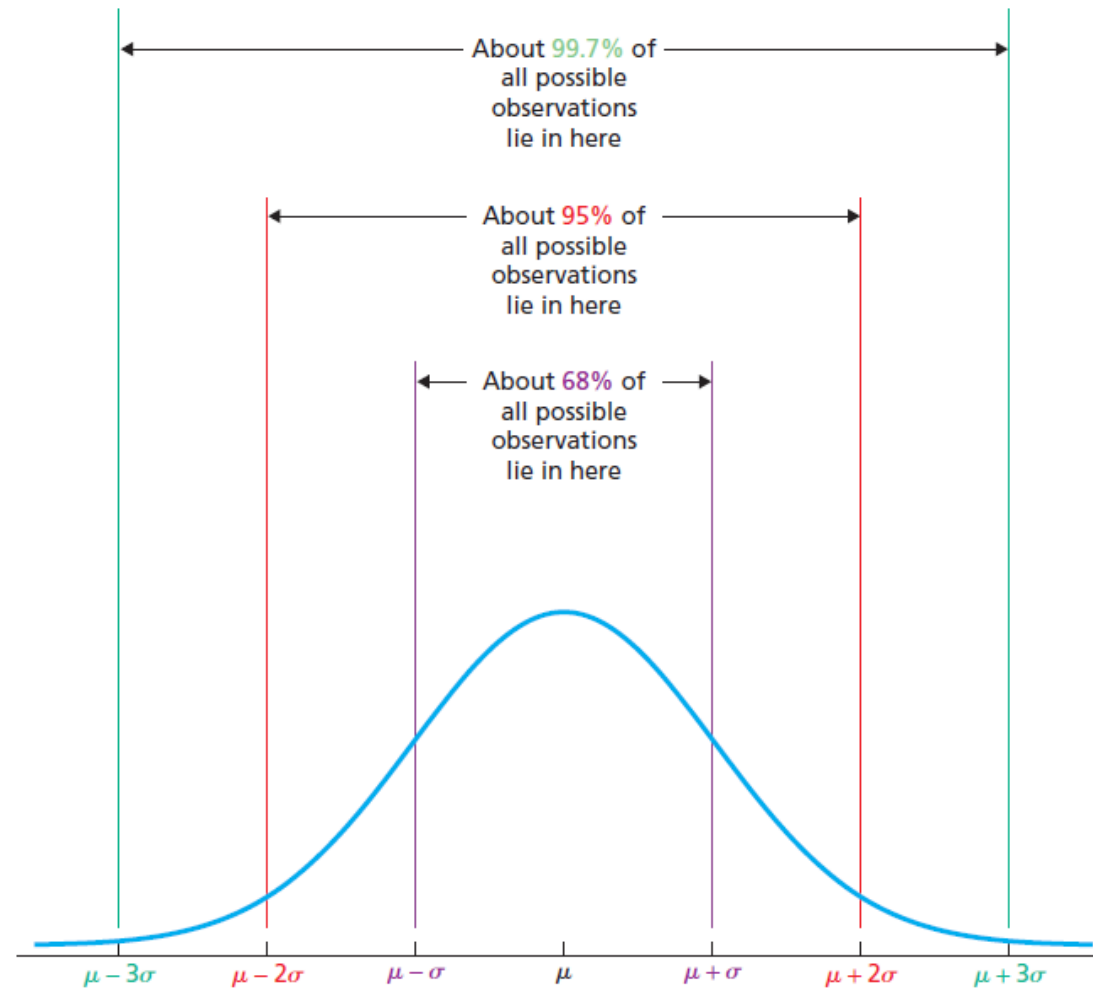
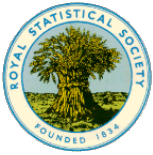
Standardized normal Curve

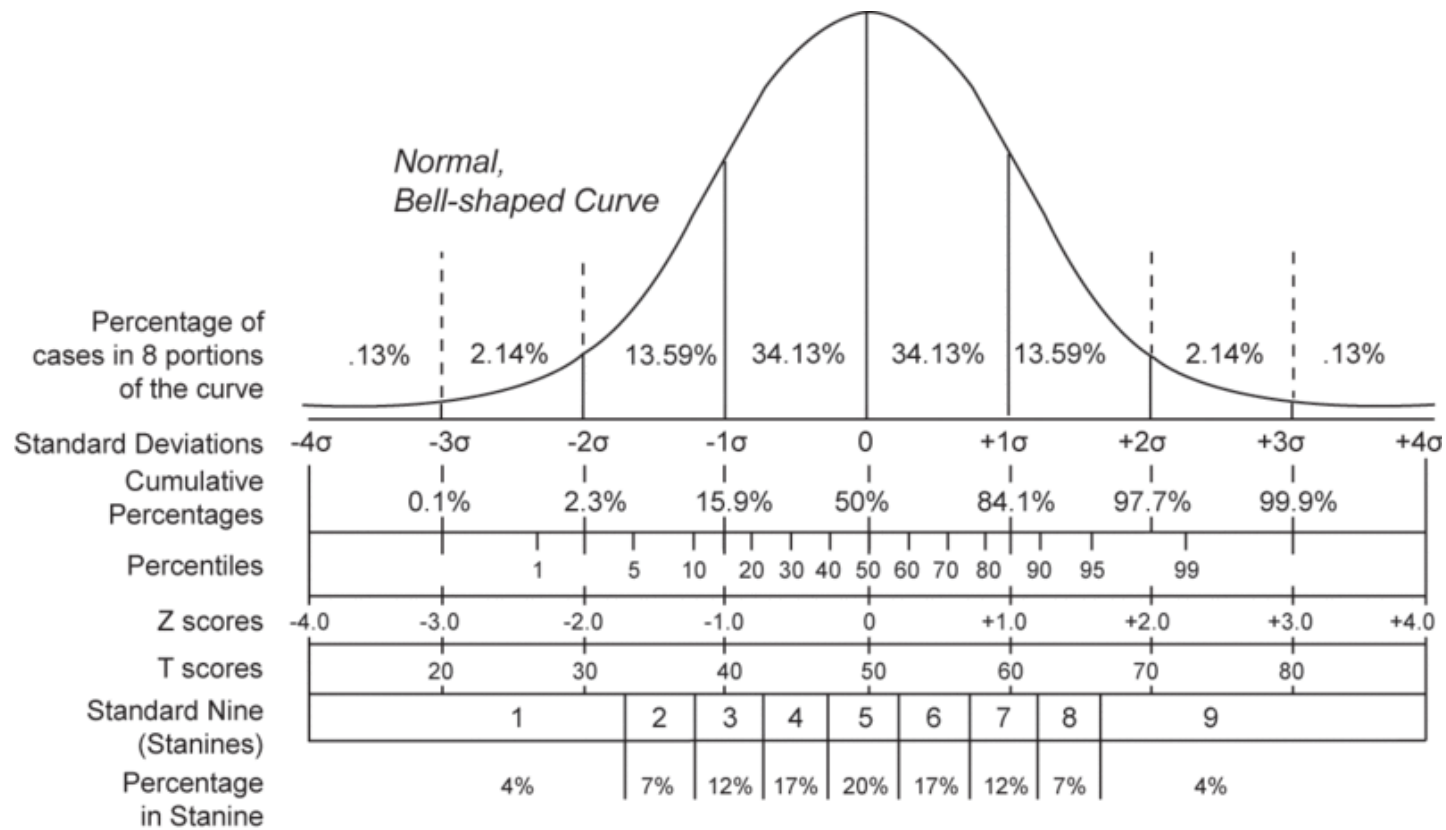
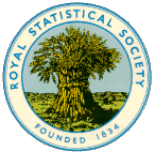
Location (mean) = 0

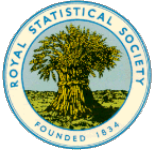
Spread (standard Deviation) = 1

Skewness = 0

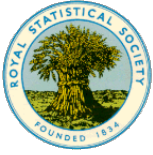
Kurtosis = 3







Assessing Normality; Normal Probability Plots

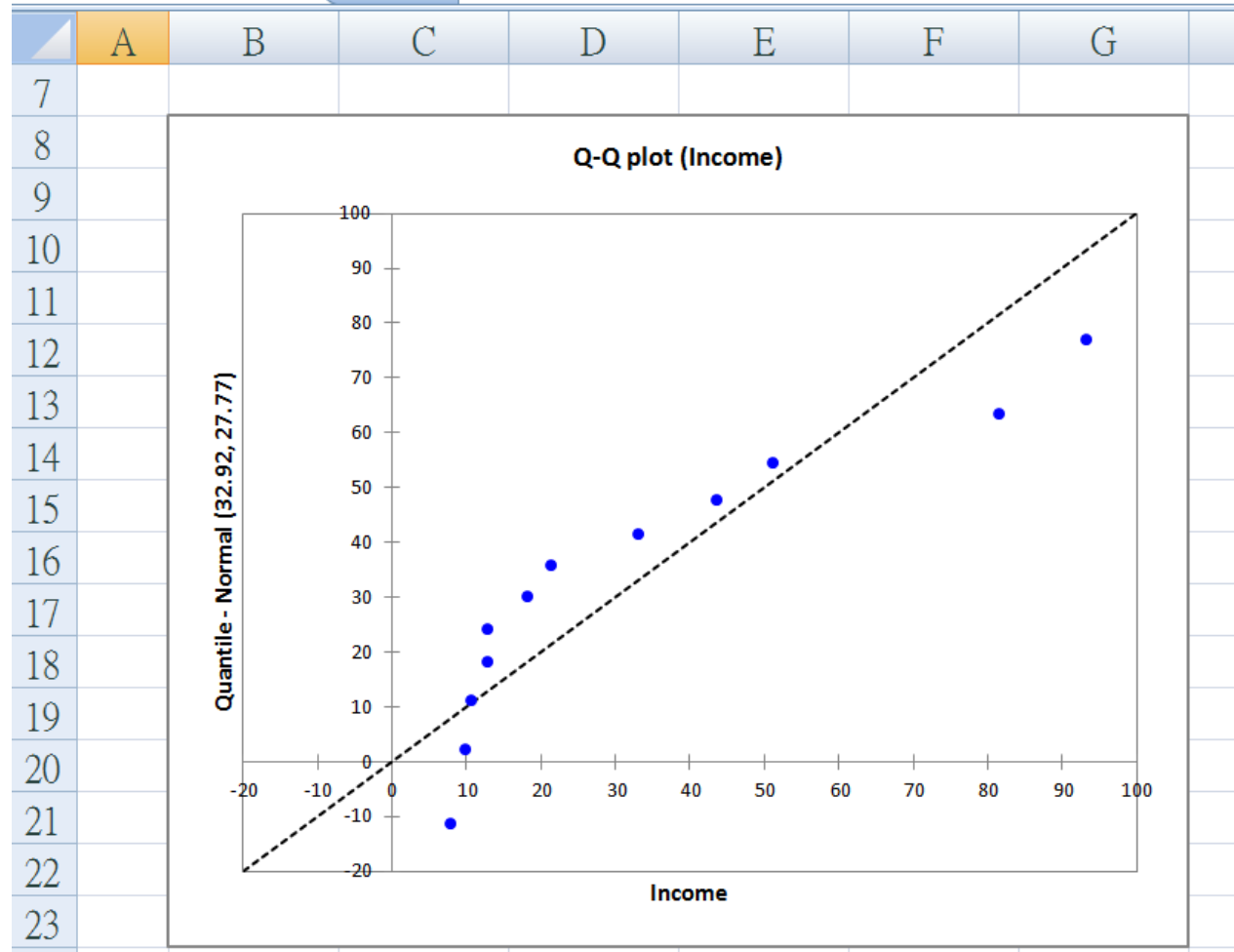
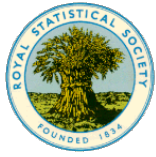


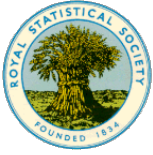
Guidelines for Assessing Normality Using a Normal Probability Plot

To assess the normality of a variable using sample data, construct a normal probability plot.

- If the plot is roughly linear, you can assume that the variable is approximately normally distributed.
- If the plot is not roughly linear, you can assume that the variable is not approximately normally distributed.

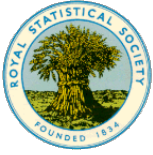
These guidelines should be interpreted loosely for small samples but usually interpreted strictly for large samples.





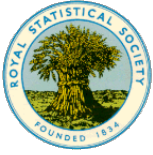
Other tests of assessing Normality

- Comparing Histogram with Normal curve
- Back-of-the-envelope test
- D'Agostino K-squared test
- Jarque-Bera test
- Anderson-Darling test
- Cramer-von Mises criterion
- Likkiefors test
- Kolmogorov-Smirnov test (STATA)
- Shapiro-Wilk test (SPSS)



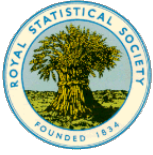
When you read a paper...

The mean age of 200 members of a social center is 42, and the median age is also 42. Is the age of 200 members is normally distributed?

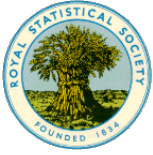


When you read a paper...

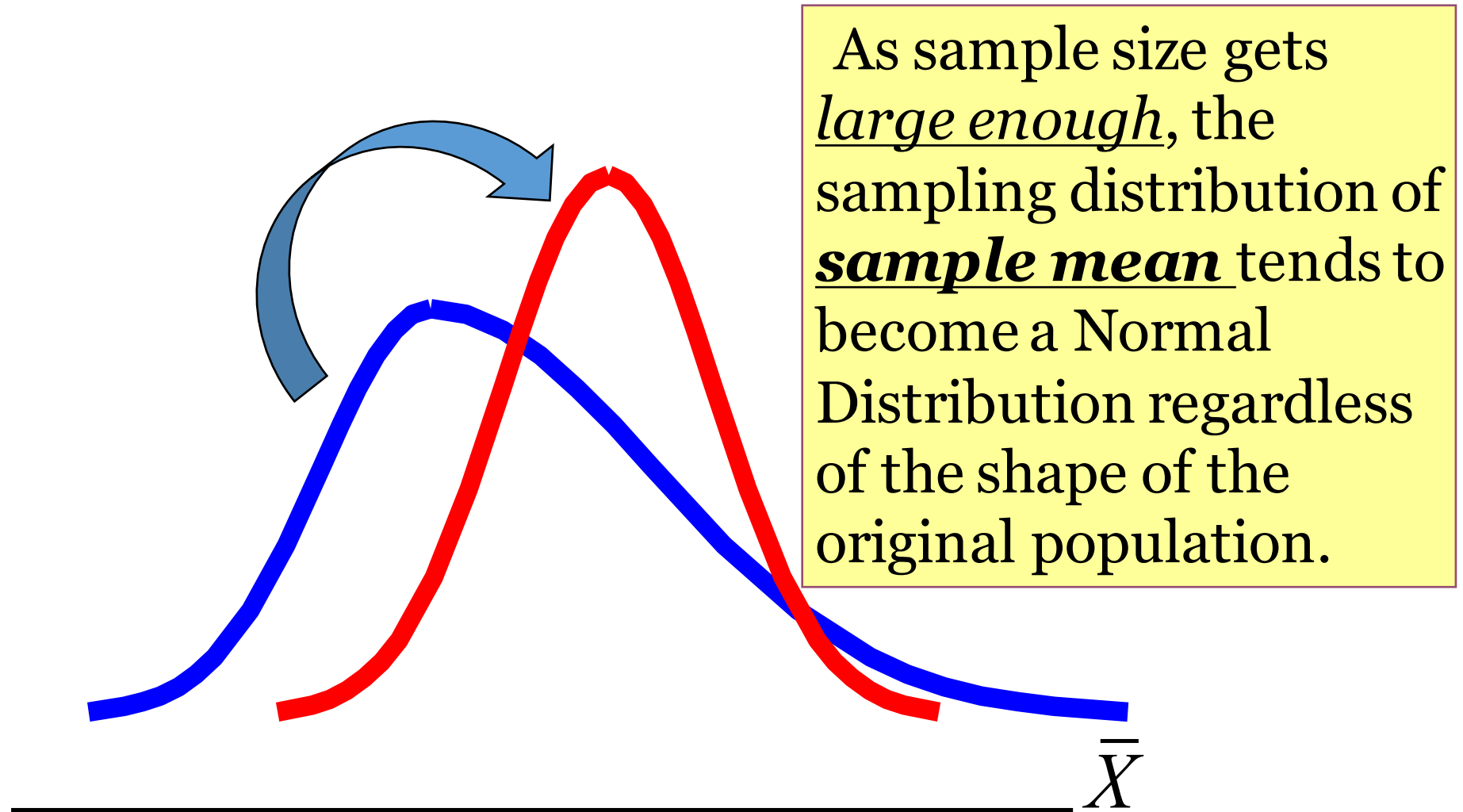
The mean score of final examination (full mark is 100) of 89 students is 75, and its standard deviation is 21. Is the score is normally distributed? If not, is it right skewed or left skewed?

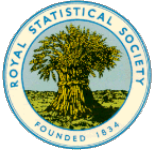


Central Limit Theorem

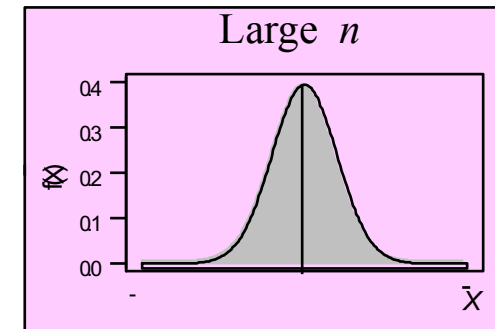
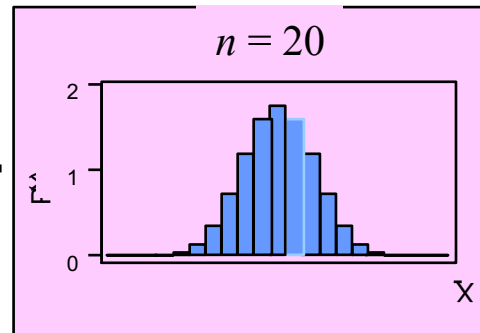
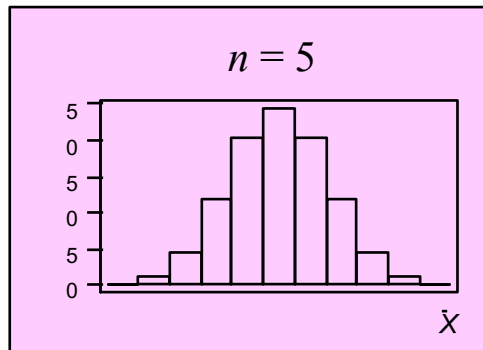


Central Limit Theorem



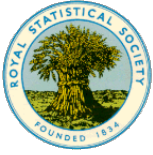


How Large is 'Large Enough' ?



- For most distributions, $n > 30$
- For fairly symmetric distributions, $n > 15$
- For normal distribution, the sampling distribution of the mean is always normally distributed
- Refer to the following website:

http://www.statisticalengineering.com/central_limit_theorem.htm

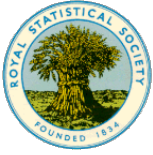


Parametric and Non-Parametric Tests

- Different definition and it is a confusing concept
 - Inferential methods concerned with parameter;
 - Methods related to theoretical distribution (t-test; chi-squared test; z-test)
 - Methods with assumption of normality of population (z-test; t-test)
- Choose appropriate tests according to data type, data distribution, sample size, ...



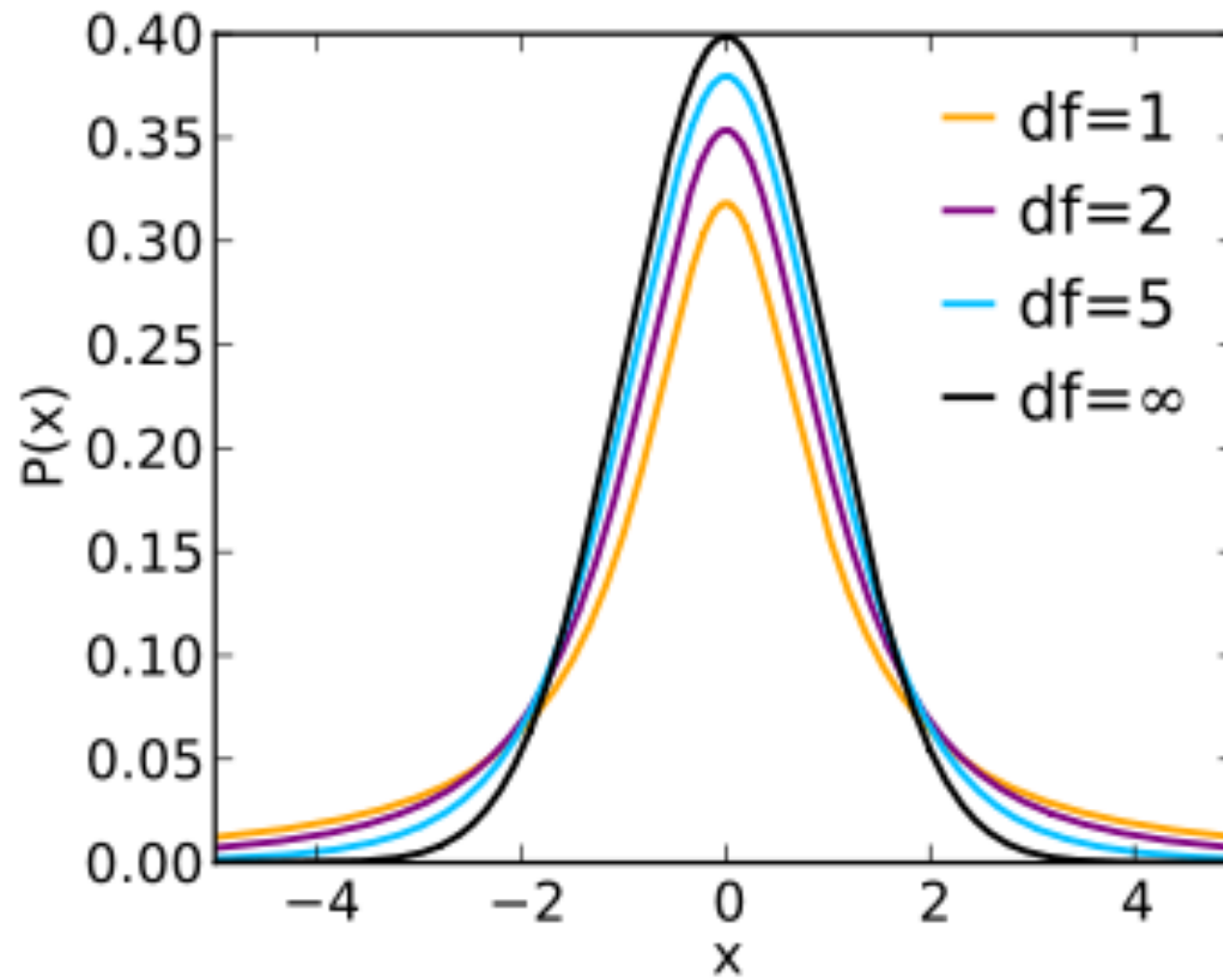
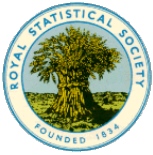
T test



History

- The t-statistic was introduced in 1908 by William Sealy Gosset, a chemist working in Dublin, Ireland ("Student" was his pen name).
- He published the test in *Biometrika* in 1908

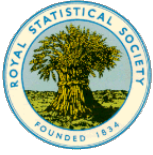






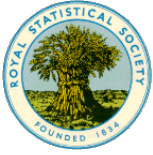
T-test for 2 independent means

- Null hypothesis
 - Two samples come from population with same means
- Assumptions of test
 - Continuous data with normal distribution
 - Variances are the same or different



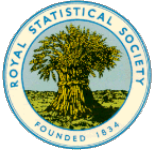
T-test for 2 independent means

- Details of the test
 - Compares means from 2 independent sample
 - Based on sampling distribution of difference of two samples
 - Allow calculate a difference and confidence interval of the difference
 - Can be calculated by formula or statistical program



T-test for 2 independent means

- If assumptions do not hold
 - The statistical test is dubious and the p value may be wrong
 - Try transformation of the data
 - It is robust to slight skewness (2 samples with same size) but is less robust if variances are clearly different
 - Skewness and different variance can be corrected by transformation.



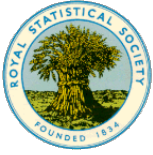
$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$$

Where

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

Degree of freedom

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}.$$



T test for paired (matched) data

- Also called one sample t-test
- It analyses mean difference in paired sample
- Null hypothesis: means difference is zero (or other values)
- Assumption
 - differences follow a normal distribution
 - variance are constant
- If assumption do not hold – transform the raw data (not the difference)



$$t = \frac{\overline{X}_D - \mu_0}{s_D / \sqrt{n}}.$$

Where \overline{X}_D and s_D is the average and standard deviation of the differences

The degree of freedom is $n-1$

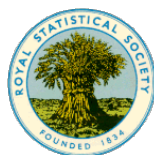
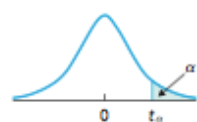
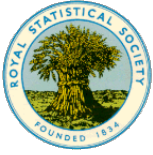


TABLE IV
Values of t_{α}



df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	df	df	$t_{0.10}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$	df
1	3.078	6.314	12.706	31.821	63.657	1	50	1.299	1.676	2.009	2.403	2.678	50
2	1.886	2.920	4.303	6.965	9.925	2	51	1.298	1.675	2.008	2.402	2.676	51
3	1.638	2.353	3.182	4.541	5.841	3	52	1.298	1.675	2.007	2.400	2.674	52
4	1.533	2.132	2.776	3.747	4.604	4	53	1.298	1.674	2.006	2.399	2.672	53
5	1.476	2.015	2.571	3.365	4.032	5	54	1.297	1.674	2.005	2.397	2.670	54
6	1.440	1.943	2.447	3.143	3.707	6	55	1.297	1.673	2.004	2.396	2.668	55
7	1.415	1.895	2.365	2.998	3.499	7	56	1.297	1.673	2.003	2.395	2.667	56
8	1.397	1.860	2.306	2.896	3.355	8	57	1.297	1.672	2.002	2.394	2.665	57
9	1.383	1.833	2.262	2.821	3.250	9	58	1.296	1.672	2.002	2.392	2.663	58
10	1.372	1.812	2.228	2.764	3.169	10	59	1.296	1.671	2.001	2.391	2.662	59
11	1.363	1.796	2.201	2.718	3.106	11	60	1.296	1.671	2.000	2.390	2.660	60
12	1.356	1.782	2.179	2.681	3.055	12	61	1.296	1.670	2.000	2.389	2.659	61
13	1.350	1.771	2.160	2.650	3.012	13	62	1.295	1.670	1.999	2.388	2.657	62
14	1.345	1.761	2.145	2.624	2.977	14	63	1.295	1.669	1.998	2.387	2.656	63
15	1.341	1.753	2.131	2.602	2.947	15	64	1.295	1.669	1.998	2.386	2.655	64
16	1.337	1.746	2.120	2.583	2.921	16	65	1.295	1.669	1.997	2.385	2.654	65
17	1.333	1.740	2.110	2.567	2.898	17	66	1.295	1.668	1.997	2.384	2.652	66
18	1.330	1.734	2.101	2.552	2.878	18	67	1.294	1.668	1.996	2.383	2.651	67
19	1.328	1.729	2.093	2.539	2.861	19	68	1.294	1.668	1.995	2.382	2.650	68
20	1.325	1.725	2.086	2.528	2.845	20	69	1.294	1.667	1.995	2.382	2.649	69
21	1.323	1.721	2.080	2.518	2.831	21	70	1.294	1.667	1.994	2.381	2.648	70
22	1.321	1.717	2.074	2.508	2.819	22	71	1.294	1.667	1.994	2.380	2.647	71
23	1.319	1.714	2.069	2.500	2.807	23	72	1.293	1.666	1.993	2.379	2.646	72
24	1.318	1.711	2.064	2.492	2.797	24	73	1.293	1.666	1.993	2.379	2.645	73
25	1.316	1.708	2.060	2.485	2.787	25	74	1.293	1.666	1.993	2.378	2.644	74
26	1.315	1.706	2.056	2.479	2.779	26	75	1.293	1.665	1.992	2.377	2.643	75
27	1.314	1.703	2.052	2.473	2.771	27	80	1.292	1.664	1.990	2.374	2.639	80
28	1.313	1.701	2.048	2.467	2.763	28	85	1.292	1.663	1.988	2.371	2.635	85
29	1.311	1.699	2.045	2.462	2.756	29	90	1.291	1.662	1.987	2.368	2.632	90
30	1.310	1.697	2.042	2.457	2.750	30	95	1.291	1.661	1.985	2.366	2.629	95
31	1.309	1.696	2.040	2.453	2.744	31	100	1.290	1.660	1.984	2.364	2.626	100
32	1.309	1.694	2.037	2.449	2.738	32	200	1.286	1.653	1.972	2.345	2.601	200
33	1.308	1.692	2.035	2.445	2.733	33	300	1.284	1.650	1.968	2.339	2.592	300
34	1.307	1.691	2.032	2.441	2.728	34	400	1.284	1.649	1.966	2.336	2.588	400
35	1.306	1.690	2.030	2.438	2.724	35	500	1.283	1.648	1.965	2.334	2.586	500
36	1.306	1.688	2.028	2.434	2.719	36	600	1.283	1.647	1.964	2.333	2.584	600
37	1.305	1.687	2.026	2.431	2.715	37	700	1.283	1.647	1.963	2.332	2.583	700
38	1.304	1.686	2.024	2.429	2.712	38	800	1.283	1.647	1.963	2.331	2.582	800
39	1.304	1.685	2.023	2.426	2.708	39	900	1.282	1.647	1.963	2.330	2.581	900
40	1.303	1.684	2.021	2.423	2.704	40	1000	1.282	1.646	1.962	2.330	2.581	1000
41	1.303	1.683	2.020	2.421	2.701	41	2000	1.282	1.646	1.961	2.328	2.578	2000
42	1.302	1.682	2.018	2.418	2.698	42							
43	1.302	1.681	2.017	2.416	2.695	43							
44	1.301	1.680	2.015	2.414	2.692	44							
45	1.301	1.679	2.014	2.412	2.690	45							
46	1.300	1.679	2.013	2.410	2.687	46							
47	1.300	1.678	2.012	2.408	2.685	47							
48	1.299	1.677	2.011	2.407	2.682	48							
49	1.299	1.677	2.010	2.405	2.680	49							

1.282	1.645	1.960	2.326	2.576
$z_{0.10}$	$z_{0.05}$	$z_{0.025}$	$z_{0.01}$	$z_{0.005}$

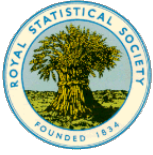


Z test



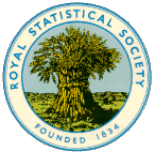
Z test

- Compares means from a sample
- Based on assumption that the sample mean distribution is normal (Central Limit Theorem for large sample size)
- Allow calculation of differences and a confidence interval for the difference
- Can be calculated by formula and statistical program



Z test

- Null hypothesis
 - The sample come from populations with the same stated mean
- Assumption of the test
 - Normal distribution of sample mean
 - Sample is large. Usually, sample size is larger than 30.



$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

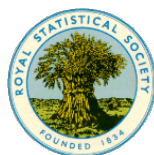
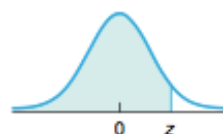
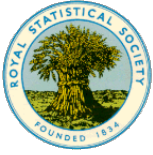


TABLE II (cont.)
Areas under the
standard normal curve



<i>z</i>	Second decimal place in <i>z</i>									
	<i>0.00</i>	<i>0.01</i>	<i>0.02</i>	<i>0.03</i>	<i>0.04</i>	<i>0.05</i>	<i>0.06</i>	<i>0.07</i>	<i>0.08</i>	<i>0.09</i>
<i>0.0</i>	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
<i>0.1</i>	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
<i>0.2</i>	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
<i>0.3</i>	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
<i>0.4</i>	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
<i>0.5</i>	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
<i>0.6</i>	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
<i>0.7</i>	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
<i>0.8</i>	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
<i>0.9</i>	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
<i>1.0</i>	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
<i>1.1</i>	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
<i>1.2</i>	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
<i>1.3</i>	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
<i>1.4</i>	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
<i>1.5</i>	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
<i>1.6</i>	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
<i>1.7</i>	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
<i>1.8</i>	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
<i>1.9</i>	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
<i>2.0</i>	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
<i>2.1</i>	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
<i>2.2</i>	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
<i>2.3</i>	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
<i>2.4</i>	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
<i>2.5</i>	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
<i>2.6</i>	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
<i>2.7</i>	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
<i>2.8</i>	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
<i>2.9</i>	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
<i>3.0</i>	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
<i>3.1</i>	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
<i>3.2</i>	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
<i>3.3</i>	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
<i>3.4</i>	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
<i>3.5</i>	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
<i>3.6</i>	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
<i>3.7</i>	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
<i>3.8</i>	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
<i>3.9</i>	1.0000 [†]									



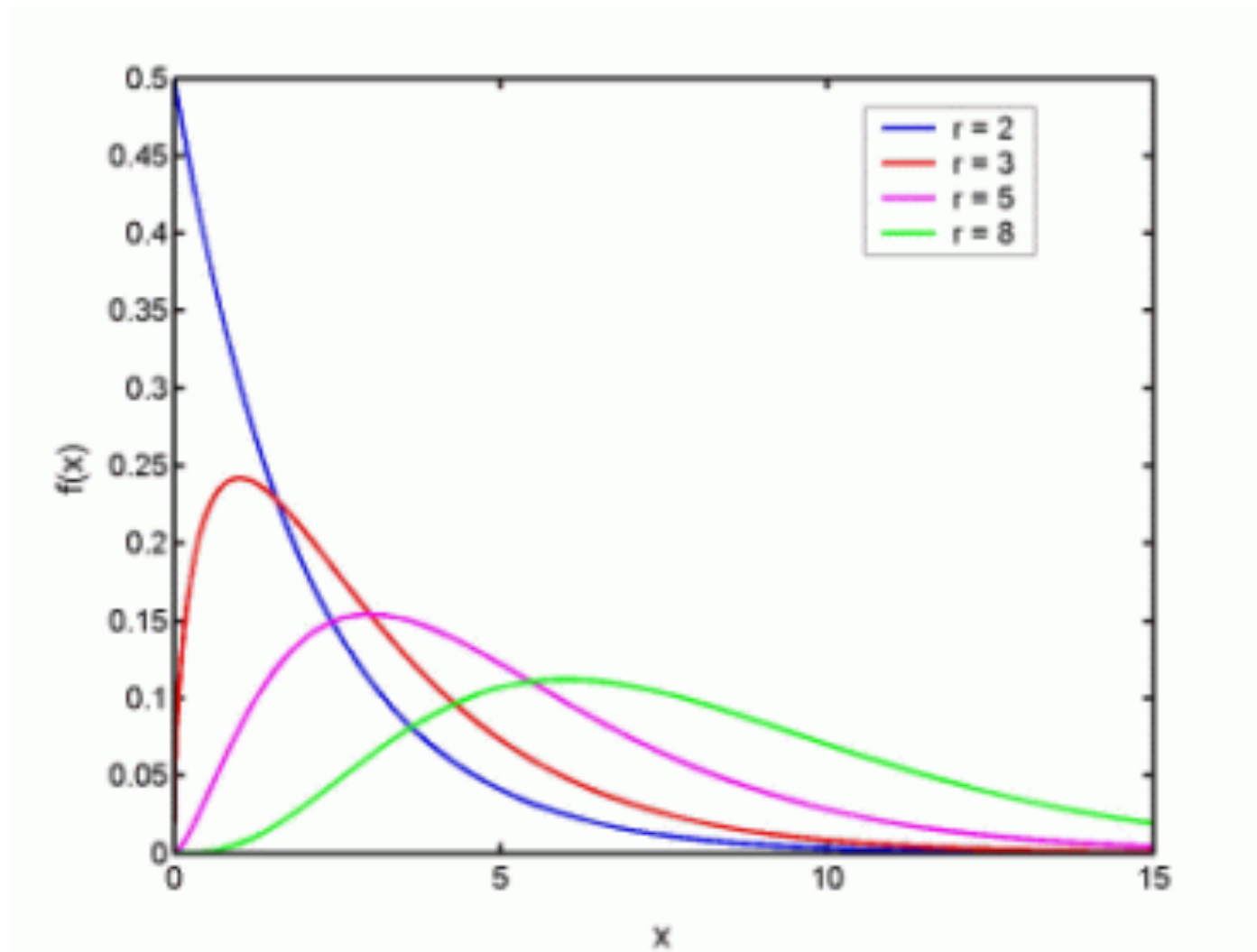
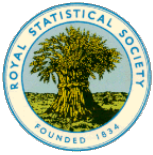
Chi squared test

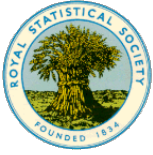


History

- **Pearson's chi-square test** is the best-known of several chi-square tests – statistical procedures whose results are evaluated by reference to the chi-square distribution.
- Its properties were first investigated by Karl Pearson in 1900.

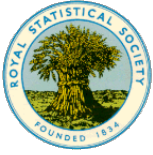






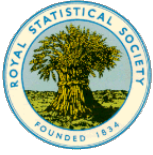
Chi-squared test

- Tests for association between two categorical variables
- Based on the chi-squared distribution with n degree of freedom
- $n = (\text{no. of row} - 1) \times (\text{no. of column} - 1)$
- It gives p value but not direct estimate or confidence interval as z test



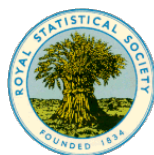
Chi-squared test

- Rationale of test
 - Calculates the frequencies that would be expected if there was no association
 - It compares the observed frequencies and expected values
 - If they are very different, this provides evidence that there is an association
 - The test uses a formula based on chi-squared distribution to give p value



Chi-squared test

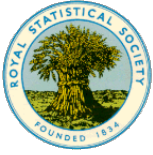
- Null hypothesis
 - There is no association between the two variables in the population from which the samples come
- Assumptions of test
 - Large sample size
 - At least 80% of expected frequencies must be greater than 5



	Gov	Prop	NP	Total
General	1680 1666.58	650 923.396	3050 2790.03	5380
Psychiatric	250 210.645	315 116.712	115 352.643	680
Chronic	20 7.435	2 4.119	2 12.446	24
Tuberculosis	5 2.478	0 1.373	3 4.149	8
Other	61 128.865	150 71.40	205 215.73	416
Total	2016	1117	3375	6508



	Gov	Prop	NP	Total
General	1680 (1666.58)	650 (923.40)	3050 (2790.03)	5380
Psychiatry	250 (210.65)	315 (116.71)	115 (352.64)	680
Chronic/TB	25 (9.913)	2 (5.492)	5 (16.59)	32
Others	61 (128.87)	150 (71.40)	205 (215.73)	416
Total	2016	1117	3375	6508



$$X^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$$

where

X^2 = Pearson's cumulative test statistic, which asymptotically approaches a χ^2 distribution.

O_i = an observed frequency;

E_i = an expected (theoretical) frequency, asserted by the null hypothesis;

n = the number of cells in the table.

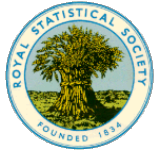
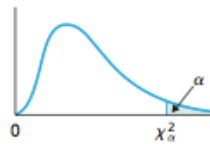
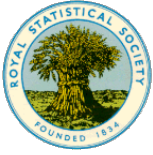


TABLE VII
Values of χ^2_{α}

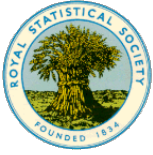


df	$\chi^2_{0.995}$	$\chi^2_{0.99}$	$\chi^2_{0.975}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.01}$	$\chi^2_{0.005}$	df
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750	5
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319	14
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	11.651	27.204	30.143	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.290	42.796	22
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559	24
25	10.520	11.524	13.120	14.611	16.473	34.382	37.653	40.647	44.314	46.928	25
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645	27
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994	28
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336	29
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672	30
40	20.707	22.164	24.433	26.509	29.051	51.805	55.759	59.342	63.691	66.767	40
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490	50
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.381	91.955	60
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.424	104.213	70
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.628	112.328	116.320	80
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.135	124.115	128.296	90
100	67.328	70.065	74.222	77.930	82.358	118.499	124.343	129.563	135.811	140.177	100



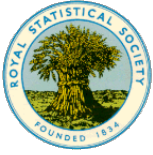
Chi-squared test

- If assumption do not hold
 - Collapsing the table
 - Continuity correction (Yates' correction)
 - Fisher's exact test
- Doing chi-squared test
 - Always use with frequencies, never use percentage
 - The formula works with all size tables
 - Can be done by computer program



Yates' Correction

- Chi-squared test based on frequencies (discrete) whilst the chi-squared distribution is continuous.
- The fit is not good in small sample size
- Yates' correction modified the chi-squared formula to make better fit
- Corrected p value (slightly bigger) should be reported



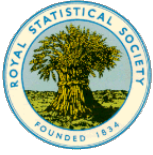
$$\chi^2_{\text{Yates}} = \sum_{i=1}^N \frac{(|O_i - E_i| - 0.5)^2}{E_i}$$

where:

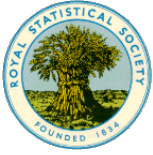
O_i = an observed frequency

E_i = an expected (theoretical) frequency,
asserted by the null hypothesis

N = number of distinct events



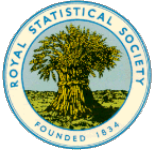
Fisher's Exact Test



History

- Fisher is said to have devised the test following a comment from Muriel Bristol, who claimed to be able to detect whether the tea or the milk was added first to her cup in 1922





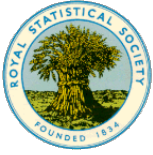
Muriel Bristol

Muriel Bristol (1888–1950)

Blanche “Muriel” Bristol was born in Croydon on April 21st, 1888 to Annie Eliza, née Davies, and Alfred Bristol, a commercial traveller. She studied to be a botanist and did a PhD on algae, probably at Birmingham University with Professor George Stephen West (1876–1919) who died in the great post-war Spanish influenza epidemic. Certainly there is a paper of July 1916 with “B. Muriel Bristol, MSc” as author with her address as the Botanical Laboratory, Birmingham University, and Bristol also wrote an obituary of West in 1921, so it seems plausible that she obtained her PhD with him. On 6 June 1923, she married William Roach (born 15 October 1895, the son of a Devon farmer) in the parish church of Saint George, Edgbaston. According to Lund, writing in 1947⁴, the species of algae, *C. muriella*, is “named after B. Muriel Bristol Roach.” Muriel Roach died in Bristol on 15 March 1950 of ovarian cancer, her age being incorrectly given on the death certificate as 62 (rather than 61). William Roach later remarried and died in 1984.



Dr Blanche Muriel Roach, nee Bristol, from a Rothamsted Research Station staff photograph. With grateful thanks to Rothamsted archivists



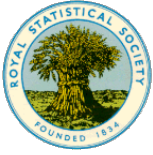
Fisher's Exact test

- Useful for small samples where chi-squared test is invalid
- Tests for an association between 2 categorical variables
- Normally used for 2 x 2 tables, but computer program allow for bigger tables
- Evaluating the probabilities associated with all possible tables which have the same row and column totals as the observed data, assuming the null hypothesis is true



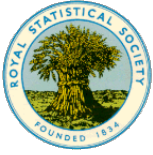
Fisher's Exact test

- Based on exact probabilities, it is computationally intensive and may be slow or fail for large sample size.
- Give p values but not direct estimate or confidence interval



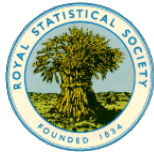
Fisher's Exact test

- Null hypothesis
 - No association between the two variables in the population from which the samples come
 - Same null hypothesis as the chi-squared test
- Assumptions of test
 - none



Fisher's Exact Test

- Always use with frequencies, never use percentages for calculation
- No simple formula, statistical program needed
- Unless with good reason, use the two-sided p value
- It gives p values at least as big as the chi-squared test. For large sample size, p values are similar
- If in doubt about the sample size, use Fisher's exact test instead of chi-squared test.



LETTER TO
THE EDITOR

Chi squared test versus Fisher's exact test

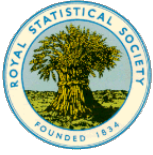
A recently published article by Chan et al¹ illustrates the problem of using the Chi squared test. If there are more than 20% of cells with an expected value of less than 5 in a contingency table, we should use Fisher's exact test. From Table 3 of their paper, we find some cells with an expected value of less than 5. For example, for the "treatment on health problem", there is a cell with an actual value of 4, its expected value is 2.09. Similar conditions occurred in the table for "Used health hotline before". In these cases, Fisher's exact test should be used instead of the Chi squared test. The former is a more accurate test, which directly calculates the probability of the distribution of the sample appearing in the table by chance. Previously, Fisher's exact test was not commonly used as the calculation procedure was tedious and complicated.² The problem has now been overcome by computers.

Interestingly, when P values in Chan et al's paper were computed, I observed that the value in Table 3 was the result obtained using Fisher's exact test and not the Chi squared test. The authors are absolutely correct to compute the P value by Fisher's exact test, but this should be stated as such instead of calling it a Chi squared test P value. The two tests are different. If in doubt about whether the sample size is large enough for the Chi squared test to be valid, use Fisher's exact test.³

KC Wong, FHKAM(Psychiatry), GradStat
Email: kc@drkcwong.com
Specialist in Psychiatry
Graduate Statistician (Royal Statistical Society)
Private practice (Psychiatrist)
Hong Kong

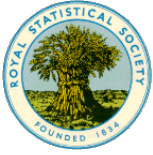
References

1. Chan FW, Wong FY, Fung H, Yeoh EK. The development of a Health Call Centre in Hong Kong: a study on the perceived needs of patients. *Hong Kong Med J* 2011;17:208-16.
2. Armitage P, Berry G, Matthews JN. *Statistical methods in medical research*. 4th ed. Massachusetts: Blackwell Publishing; 2002: 134-7.
3. Peacock JL, Peacock PJ. *Oxford handbook of medical statistics*. Oxford: Oxford University Press; 2011: 266.

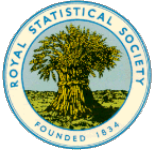


a	b	a + b
c	d	c + d
a + c	b + d	a + b + c + d

$$p = \frac{\binom{a+b}{a} \binom{c+d}{c}}{\binom{n}{a+c}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{a!b!c!d!n!}$$

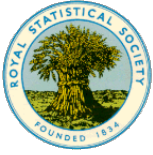


Mann Whitney U test (Wilcoxon two sample signed rank test)



History

- It is also called as Wilcoxon two sample signed ranked test.
- It was proposed initially by Frank Wilcoxon in 1945, for equal sample sizes, and extended to arbitrary sample sizes and in other ways by Henry Mann and his student Donald Ransom Whitney in 1947



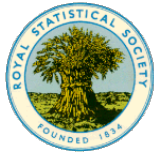
Mann-Whitney U test

- It is based on the rank of the data
- It gives p value but no estimate
- Given a table of cut-offs, the test is easy to do by hand
- If the sample is very small (both smaller than three observations), then statistical significant is impossible.

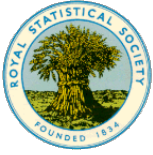


Mann-Whitney U test

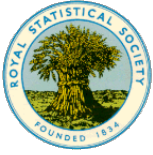
- The samples should be unrelated (independent)
- The test can be applied to ordinal data
- The population distributions should have the same shape (it tests the whole distribution)
- Tied values in the data
 - If number of ties is small, it still satisfactory



n_2	α	n_1							
		3	4	5	6	7	8	9	10
3	0.10	14	20	27	36	45	55	66	78
	0.05	15	21	29	37	46	57	68	80
	0.025	—	22	30	38	48	58	70	82
	0.01	—	—	—	39	49	59	71	83
	0.005	—	—	—	—	—	60	72	85
4	0.10	16	23	31	40	49	60	72	85
	0.05	17	24	32	41	51	62	74	87
	0.025	18	25	33	43	53	64	76	89
	0.01	—	26	35	44	54	65	78	91
	0.005	—	—	—	45	55	66	79	93
5	0.10	18	26	34	44	54	65	78	91
	0.05	20	27	36	46	56	68	80	94
	0.025	21	28	37	47	58	70	83	96
	0.01	—	30	39	49	60	72	85	99
	0.005	—	—	40	50	61	73	86	101
6	0.10	21	29	38	48	59	71	84	98
	0.05	22	30	40	50	61	73	87	101
	0.025	23	32	41	52	63	76	89	103
	0.01	24	33	43	54	65	78	92	106
	0.005	—	34	44	55	67	80	94	108
7	0.10	23	31	41	52	63	76	89	104
	0.05	24	33	43	54	66	79	93	107
	0.025	26	35	45	56	68	81	95	110
	0.01	27	36	47	58	71	84	98	114
	0.005	—	37	48	60	72	86	101	116
8	0.10	25	34	44	56	68	81	95	110
	0.05	27	36	47	58	71	84	99	114
	0.025	28	38	49	61	73	87	102	117
	0.01	29	39	51	63	76	90	105	121
	0.005	30	40	52	65	78	92	108	124
9	0.10	27	37	48	60	72	86	101	116
	0.05	29	39	50	63	76	90	105	121
	0.025	31	41	53	65	78	93	108	124
	0.01	32	43	55	68	81	96	112	129
	0.005	33	44	56	70	84	99	114	131
10	0.10	29	40	51	64	77	91	106	123
	0.05	31	42	54	67	80	95	111	127
	0.025	33	44	56	69	83	98	114	131
	0.01	34	46	59	72	87	102	119	136
	0.005	36	48	61	74	89	105	121	139

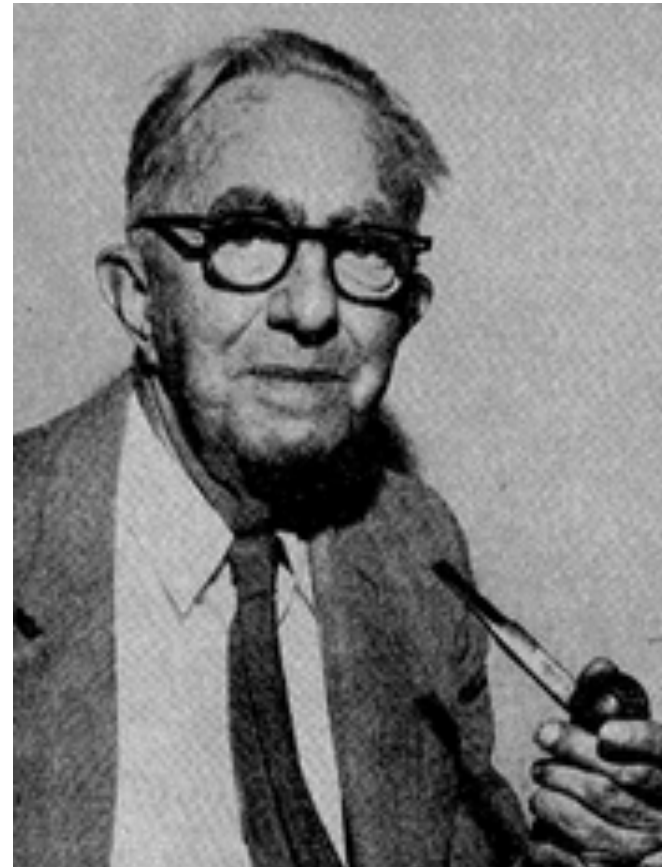


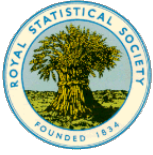
Wilcoxon matched pairs test



History

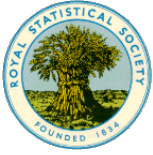
- It also called Wilcoxon signed rank test
- The test is named for Frank Wilcoxon (1892–1965) who, in a single paper, proposed both it and the rank-sum test for two independent samples (Wilcoxon, 1945)





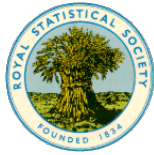
Wilcoxon matched pairs test

- Based on the signs of the differences in the pairs and the relative sizes of differences rather than the actual values
- It gives p values but no estimate
- Given a table of cut-offs, the test is easy to do by hand
- If the sample is smaller than 7, then statistically significance is impossible

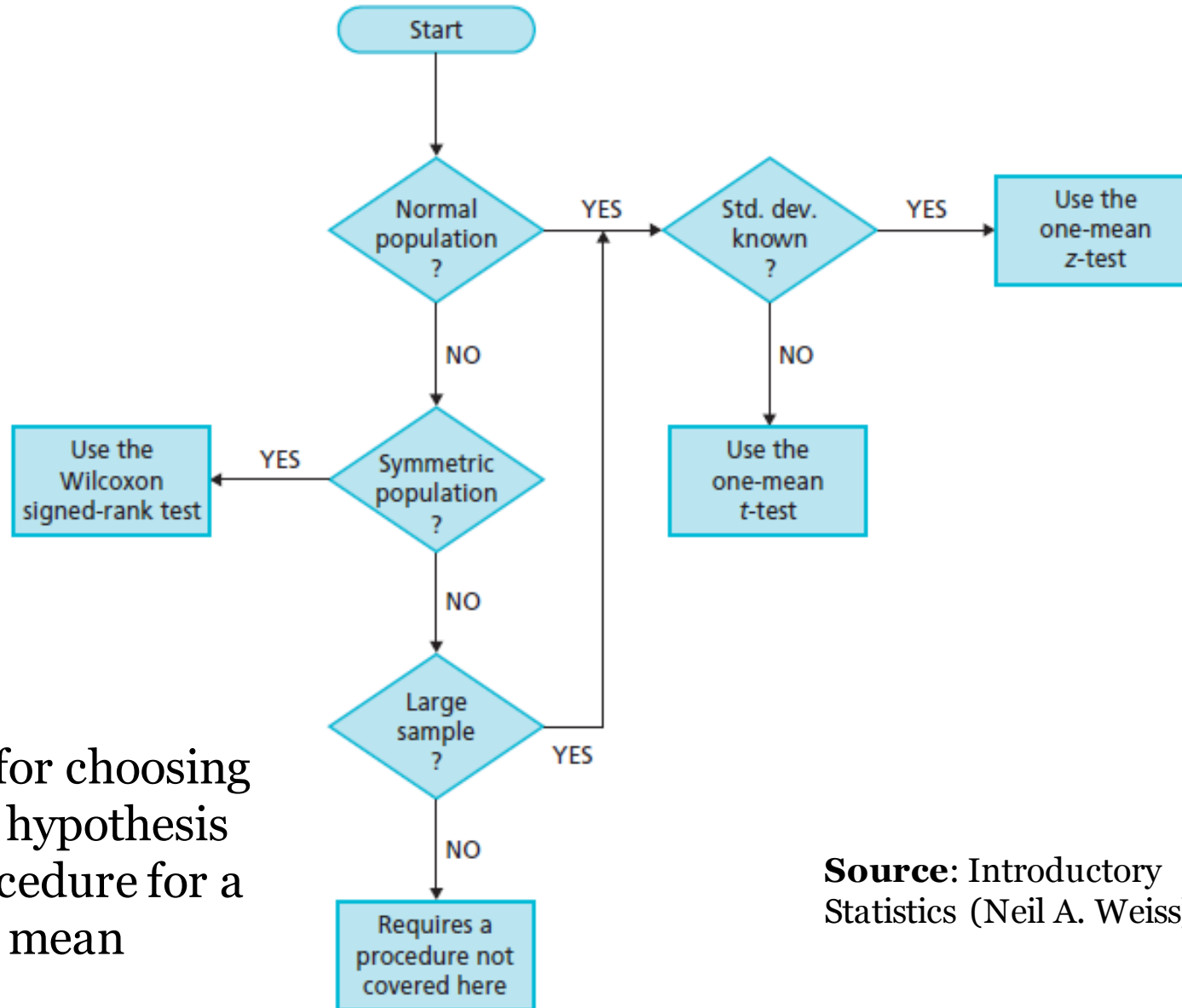
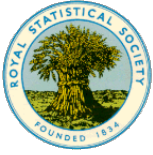


Wilcoxon matched pairs test

- The data should be paired
- The data should consist of numerical measurements (interval data)
- The distribution of the differences should be symmetrical. It is difficult to know in practice whether the distribution is symmetrical or not, but we can apply it with small sample size
- The zero difference (tie) should be omitted

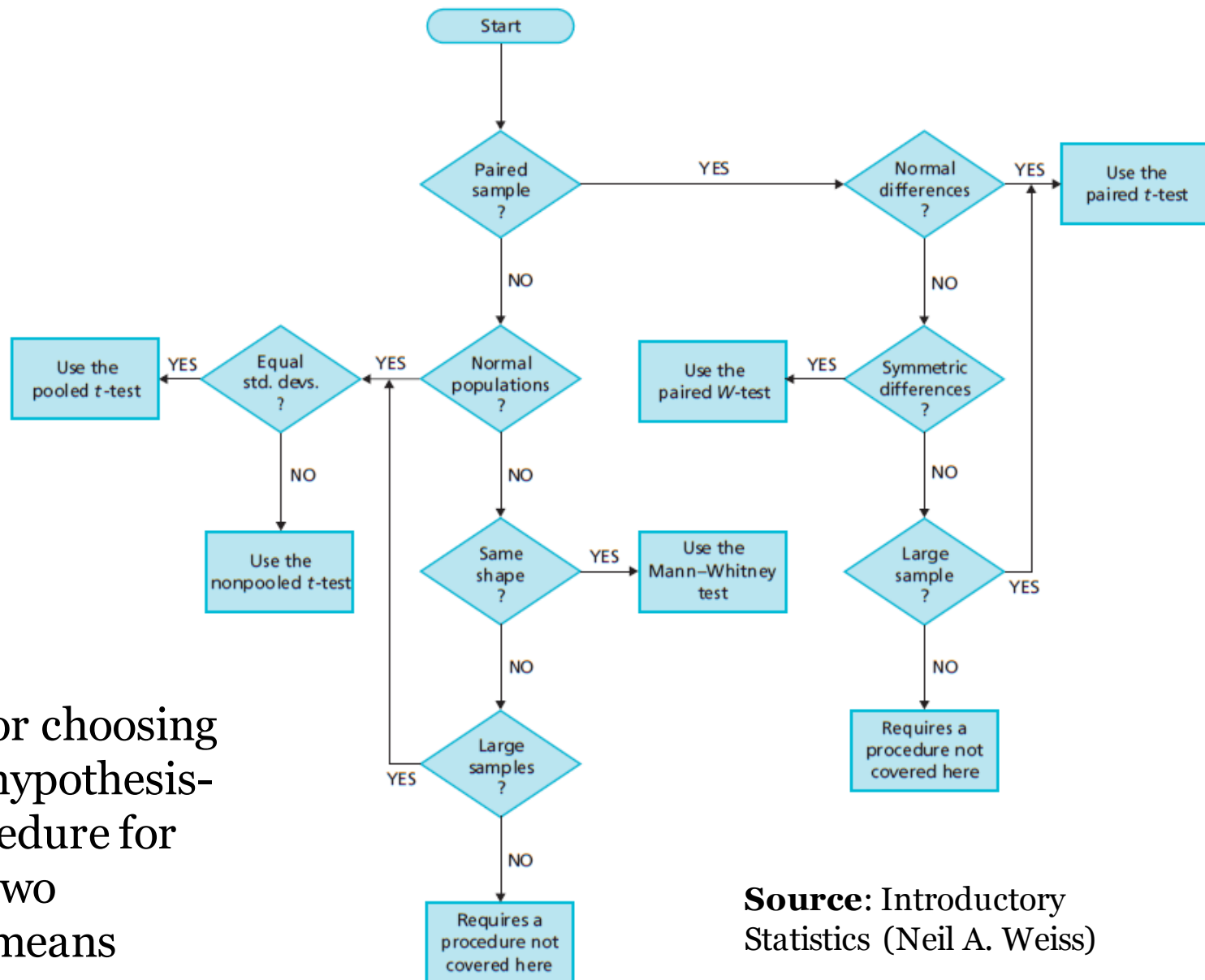


<i>n</i>	$W_{0.10}$	$W_{0.05}$	$W_{0.025}$	$W_{0.01}$	$W_{0.005}$	<i>n</i>
7	22	24	26	28	—	7
8	28	30	32	34	36	8
9	34	37	39	42	43	9
10	41	44	47	50	52	10
11	48	52	55	59	61	11
12	56	61	64	68	71	12
13	65	70	74	78	81	13
14	74	79	84	89	92	14
15	83	90	95	100	104	15
16	94	100	106	112	117	16
17	104	112	118	125	130	17
18	116	124	131	138	143	18
19	128	136	144	152	158	19
20	140	150	158	167	173	20



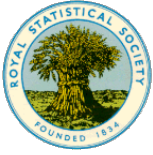
Flowchart for choosing the correct hypothesis testing procedure for a population mean

Source: Introductory Statistics (Neil A. Weiss)



Flowchart for choosing the correct hypothesis-testing procedure for comparing two population means

Source: Introductory Statistics (Neil A. Weiss)



**THANK YOU
FOR
YOUR
ATTENTION!
ANY QUESTIONS?**